DOCUMENT RESUME

ED 426 331 CG 029 015

AUTHOR	Herr, Edwin L.
TITLE	Assessment '98: A Hotbed of Issues and Challenges.
PUB DATE	1998-01-00
NOTE	34p.; Paper presented at the ERIC Clearinghouse on Counseling and Student Services Conference: "Assessment '98: Assessment for Change - Changes in Assessment" (St. Petersburg, FL, January 16-18, 1998).
PUB TYPE	Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Accountability; Change; Counseling; Counselor Role; Cultural Pluralism; Educational Policy; Elementary Secondary Education; Employment Practices; Employment Qualifications; *Evaluation; *Politics of Education; *Public Policy; School Counseling; School Counselors; Test Bias; Test Interpretation; *Test Use; Testing; Validity
IDENTIFIERS	*Diversity (Groups)

ABSTRACT

The status of assessment in terms of federal policy, the social or economic climate related to education, the role of testing in counseling, assessment as an intervention in its own right, and assessment as it is affected by technology are important components of the context of issues and challenges related to assessment at the edge of the 21st century. This lecture raises questions about assessment and its uses in education, the workplace, and society. Because bias is intrinsic to the form of assessment used, the problem cannot be eliminated simply by changing the test questions. The sociopolitical context as well as the psychometric properties of tests, the use of tests for accountability, the relationship of standards to the means provided for meeting those standards, and the several validities that any measurement procedure has, must be considered. Questions are raised on issues surrounding infrastructure, changing social values, the political nature of the debate, technical capacity, national standards, training, response to media and policy makers, and purposes of testing by school counselors. Several issues such as those concerning testing and licensing between the several professional organizations involved and issues of "teaching the test," computer applications, ethics and the Internet, and training counselors, are discussed. (EMK)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 426 331

Assessment '98: A Hotbed of Issues and Challenges

Edwin L. Herr
The Pennsylvania State University

It is a great privilege to participate in this conference, "Assessment '98: Assessment for Change – Changes in Assessment" and to address specifically "Assessment '98: A Hotbed of Issues and Challenges". Although I did not suggest the term "hotbed of issues and challenges" as a title for my remarks it is certainly a concept which is "apropos" as we look at the status of assessment in terms of federal policy, the social or economic climate related to education, the role of testing in counseling, assessment as an intervention in its own right, or assessment as it is affected by technology. Such emphases, while not exhaustive, are important parts of the context in which the issues and challenges related to assessment can be framed at the edge of the 21st Century and as we peer over that artificial time boundary of the year 2000 into the next several decades. Given the lack of absolute certainty we have about projections related to assessment in the years ahead, we may find it useful in this conference to raise a series of questions about assessment that will help us, as professionals, to engage in long-term strategic planning as a way of reducing the uncertainty we experience as we consider the likely changes ahead. I will try to sprinkle some of these questions through my remarks.

As a final prefatory comment, let me say that I have been asked to speak today not as an assessment expert, which I am not, but rather as a long-time educational administrator and counselor who has been and is concerned about the results of assessment and the issues that affect such processes.

Given these preliminary perspectives, let me try to construct an "external" view of assessment and some of the issues and challenges that I see embedded in such a view. First, I

CG029015

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
E. HERR
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

BEST COPY AVAILABLE

would argue that at the present time in the United States, assessment is being expected to perform roles and functions that are unprecedented in its history. As Terwilliger (1997) has recently observed, "It is obvious to even the most casual reader of the literature on educational assessment that the field is currently undergoing a fundamental and profound transformation. The traditional concepts and methodologies associated with assessment are being questioned by a variety of critics including school reform advocates, subject matter experts, cognitive theorists, and others. In general, advocates for change recommend that at assessments of achievement should be designed to reflect more precisely complex 'real-life' performances and problems than is possible with short-answer and choice-response questions that characterize many teacher-made tests" (p.24).

Whether justified or not, implications for assessment pervade the language of federal and state policy directed to the health care industry, to manufacturing and financial processes, to both basic and higher education, and to the allocation of resources. Whether you are in a University or a school district, in community mental health services, or in other social and economic institutions, terms like data-driven, standards, performance indicators, continuous quality improvement, total quality management, accountability, benchmarking, strategic initiatives and strategic actions, competency, certification, accreditation, licensure have become standard vocabulary and operating processes which define much of our professional existence. Each of these terms, as they are implemented, embody some form of assessment, measurement, or testing. This reality has promoted one anonymous wag to suggest that in the constitution of the United States we have replaced the creed of the founding fathers that "We hold these truths to be self-evident" with the words "We hold these truths to be statistical."

Although assessment in the United States is still young in chronological terms, dating, in operational terms, depending on your assumptions, only to the first two decades of the twentieth century when Binet brought his intelligence test from France to the United States, or the Army Alpha tests were developed for use in World War I, or the first vocational interest measure was given preliminary attention at the first Applied Psychology department in the nation at what was then called Carnegie Tech in Pittsburgh in 1915, assessment in the United States has achieved much and grown as a science during the last seven or eight decades. But throughout the twentieth century, assessment has tended to be used for rather specific purposes and within restricted contexts. Even though there are exceptions to this point, assessment has not been consistently defined by government policy as a sociopolitical instrument of national importance. Although the National Defense Education Act, the Armed Forces, and the federal enabling legislation for employment counselors and rehabilitation counselors have certainly emphasized testing to identify either gifted adolescents who should be encouraged to enter science and mathematics in higher education, or to identify the performance capabilities of inductees into the military or of the unemployed or of persons with disabilities for whom specific training should be provided, in general assessment purposes and uses have evolved incrementally as assessment knowledge and techniques have evolved, their purposes and uses, in most cases, have been limited, and they have not been the focus of federal policy debate.

At the moment, however, there appears to be a national love-hate climate surrounding assessment as such processes have become partisan political grist for public policy debates between the Republican Party and the President of the United States or between political parties and special interest groups. Examples of these issues lie with the recent headlines about the President's commitments to national academic standards in reading and mathematics and

voluntary national testing to determine whether these standards are being met versus Representative William Goodling, Chair of the House Education and Workplace Committee, who does not think federal money should be focused on testing but on better education (Hoff, 1997). In addition, Representative Goodling's plan in relation to that of the President's would require the National Academy of Sciences to review all exiting commercial tests to determine if they can create "an equivalency scale" to compare students' scores on them". Parallel to Representative Goodling's request, President Clinton has assigned the Governing Board of the National Assessment of Educational Progress to study his proposal for new national tests, their use and their design. (Lawton, 1997, December 3) While there have been recent compromises between the parties on the President's initiatives, related to voluntary testing, other federal and state legislation continues to elevate assessment as a major strategy. For example, Massachusetts (White, 1997) will require by 1999-2000 statewide learning standards and assessment in core subjects. The Chief Education Officer in Massachusetts is currently arguing for a mandatory high school graduation examination such as the GED. New Jersey is trying to link funding levels for schools to statewide academic standards and their assessment (Johnston, 1997). Rhode Island (Archer, 1997) is now using criterion referenced tests in selected academic subjects to measure how students do when compared against a state goal for performance rather than the previously used Metropolitan Achievement Tests. In the new criterion referenced tests they are using, the State has defined how good is good enough rather than using national norms on specific standardized instruments. Rhode Island's approach to assessment, particularly in math achievement, is related to the efforts of the New Standards Project, a collaborative effort of more than a dozen states that developed standards and related assessments for student performance as

these standards are benchmarked against national and international standards of what students should know and should be able to do.

Texas (Lawton, 1997) has developed the Texas Assessment of Academic Skills that uses standardized tests as gate keepers to high school graduation. North Carolina (Manzo, 1997) has intensified its focus on teaching a state curriculum around which assessments were designed to hold schools and school districts accountable for student achievement. Michigan (Johnson, 1997) has passed a bill to revise their high school testing program to grant "state endorsements" in math, science, social studies and communication arts. The endorsements that would be graded by student performance level—basic, above average, or outstanding—would appear on transcripts instead of on diplomas.

IDEA (the Individuals with Disabilities Education Act) (Sack, 1997) has proposed new rules that provide new requirements for an individualized education plan that each disabled student must have and the inclusion of disabled students in academic assessments. These rules require that all students with disabilities must be included in state or district assessments, or be given an alternative examination. Further, these roles require that states must set performance standards, similar to those for non-disabled students, for students receiving special education services.

As you are well aware, other states and federal policies could be identified here to illustrate the comprehensiveness with which academic standards, accountability and assessment are now seen as interactive by federal and state policy. But it also needs to be noted that such interactions are not moving forward without challenge. For example, the Texas Exit Exam has been the focus of lawsuits because the passage rates of African-Americans and Latinos in that State are significantly below that of White students. The North Carolina assessments are being

challenged by parents and some testing experts who argue that the North Carolina end of grade tests were designed to hold schools and school districts accountable but they are being used for individual assessment to determine whether students would be held back or promoted or need remediation. The Michigan tests are the focus of parental objections because they are too time-consuming (they have taken nine hours but are being shortened to six hours) and they unfairly label students who do poorly on tests.

The National Association of State Boards of Education (Lawton, 1997) has contended that assessments must be in tune with rigorous state standards, address specific goals, offer some national and international comparisons, include all students, and be thoroughly evaluated . . . In this view, an effective assessment system should also help a state identify learning groups and high achievement . . . the state then has the obligation to follow up, providing help to the students who still need to meet academic goals or to guide the offering of more instruction to foster continued achievement in the most accomplished of students. But, the National Association of State Boards of Education also provides the following cavea:

"However, denying a diploma based only on test scores when the student is otherwise qualified to graduate means that students who do well in school but perform poorly on the State assessment may be unfairly penalized by a one-shot evaluation of their accumulated school work."

However you define these issues and challenges, assessment has become a high stakes mechanism affecting the life chances of many young people and substantially defining the curriculum to which teachers will teach in order to have their students perform as well as possible in State assessments to which they will be exposed. Currently, some 26 states rely

entirely or nearly so, on multiple choice tests to measure student knowledge and skills in all subjects.

Obviously, embedded in these trends at federal and state levels are a hotbed of issues and challenges that either revisit continuing and recurring questions or identify emerging and future challenges. These include issues that focus on the tests as emblematic of potential federal control over all subject matters, of a potential national and centralized curriculum, of the inappropriateness of the proposed tests to identify individual deficits in the academic areas being assessed, and the use of national tests to be used only in English to measure reading and mathematics of children whose first language is Spanish or Chinese as is true in specific school districts around the country (for example, California and Texas).

Among the major issues that continue to ferment are explicit or implicit concerns about test bias and gender differences. One of the new perspectives on these issues come from studies from Supovitz (1997) and his colleagues which contend that standardized tests using a multiple choice format are the predominant form of assessing the achievement of America's children. Supovitz argues that a diverse society deserves a more diverse assessment system. He contends that "of course standardized tests are biased. But it is not just standardized tests—any simple testing method is biased because it applies just one approach to getting at student knowledge and achievement. Any single testing method has its own particular set of blinders. Since the bias in testing is intrinsic in the form of assessment used, we cannot eliminate this problem simply by changing the questions asked. Rather we must ask the questions in many different ways. "(p.34) . . . today's large-scale, largely multiple-choice assessments exist in a vacuum. They stand alone, inflating their importance. Since there are no forms of assessment that, in combination with standardized tests can provide a more robust image of a student's capabilities, we have

come to rely on one particular type of assessment as the measure of student achievement. Standardized tests are the only game in town" . . . "What we need are more experiments employing combinations of assessment approaches to arrive at an appropriate melding of test forms both economically feasible and robust enough to minimize the bias inherent in any single measure alone" . . . "In the end, the larger more intractable sources of disparities in student performance stem from broad social and educational inequities. But within the realm of assessment, the challenge for educators and policy makers is to find the appropriate balance of a variety of assessment forms, so that students of different genders, from different backgrounds, and with different affinities can demonstrate their capabilities" (p.37).

The perspectives of Supovitz about test bias leads to some related perspectives that are inherent if not explicit in the debates about specific uses of assessment by parents, minority groups, politicians and, indeed, testing experts. Coming from a multicultural perspective is the continuing concern of many observers that testing is sexually or racially biased and, indeed, penalizes rather than facilitates the growth of specific groups of clients. While some of the recommendations of Supovitz would be helpful in ameliorating such matters, still others argue that the reasons for testing during this century have changed and that purposes and uses of testing and assessment must change accordingly. As I read of the debates in Washington and in the states about assessment, I am frequently reminded of the important insights of Gordon and Terrel (1981) a decade and a half ago. They stated:

> "Critics of testing argue from a sociopolitical context, and thus challenge the very purpose as well as the developed technology of standardized testing. Defenders of testing argue from a traditional psychometric context, with little or no concern for political and

social issues. The arguments of the two parties cannot be understood and appreciated without reference to those contexts (p.1167)."

Gordon and Terrell argued that the reasons for testing at the beginning of the twentieth century and several decades afterward have changed and so must the purposes of testing. The meritocratic selection of a few as a goal has given way to a shift in the approach to allocation of opportunities in response to changes in the social and political environment. The assertions of a group superiority on the basis of test scores and the subsequent control of the opportunity and reward structure to retain low-status groups in some socially assigned position has given way to an attempt to democratize access to opportunity; thus the use of tests also should change. As understanding grows about the pluralism in and diversity of the effects of ethnicity, sex, race, and social class upon cognitive and affective structures, learning styles, motivation and related matters, these should be reflected in purposes for assessment. In a sort of precursor to the current perspectives of the Individuals with Disabilities Education Act about testing, to again quote Gordon and Terrell:

> "The proper course of assessment in the present age is not merely to categorize an individual in terms of current functioning, but also to describe the processes by which learning faculty and disability proceed in a given individual so that it is possible to prescribe developmental treatment if necessary . . . The equalization of opportunity may require that intervention be responsive to the functional characteristics of the person to whom the opportunity is being made available. It must be determined where the examinee is in terms of function, how he or she got there, and how growth within the examinee's particular social and cultural environment can be enhanced" (p.1170).

Former New York Governor, Mario Cuomo, in a recent speech to the Council of the great City Schools' annual conference, made the above points succinctly when in relation to standards and assessment, "If we're going to set the bar high, we're going to have to have all the things we need to get the children over that bar" (Reinhard, 1997). Such an issue is not only relevant to children put to adolescents and adults as well.

Underlying the perennial debates about assessment and the use of tests for accountability and the other purposes mentioned, but not always well articulated is the reality that any test, assessment, or other measurement procedure has many validities, not one (Messick, 1995). In fact, it is not only the validity of the measure itself about which researchers, policy-makers, teachers or counselors must be concerned but rather the validities of the inferences from the measures that are made. Thus, however scientific or empirical the development of any measurement instrument may be, its probable multiple validities and the inferences that can be made from it bring both the test and the inferences into the area of values and social contexts. Frequently, then, those who argue for or against tests are really arguing about the different validities or inferences that can be assigned to these tests. Thus, whether or not we know it, many of the controversies about standardized tests and other forms of assessment can be dismantled into issues which have to do with the constructive or predictive validity of tests on the one hand, and such issues as the utility of test information, or perhaps more precisely, the social functions of standardized tests on the other.

While I have lingered on current federal policy in education as a hotbed for questions and challenges to assessment, many of the same concerns apply to the use of testing and assessment in other contexts, universities and workplaces, and also raise additional questions of particular

relevance to counselors. Before I turn to these issues briefly, let me raise a number of questions which evolve from the observations I have made so far.

1. How do we develop the infrastructure to create the assessment strategies necessary for a diverse society that merits the resources and time to invent, test, and integrate into a program of assessment the measures—performance, portfolio, etc.—that go beyond standardized, multiple choice instruments?
2. Are we fully aware that the use of assessments do not exist in a vacuum; their purposes are shaped by changing social values as well as by their psychometric properties?
3. Are we as a profession of persons interested in the development of or the use of assessment, fully attentive to the reality that much of the debate about tests is political, not scientific? Do we have the will and the insights to enter that debate and bridge the often disparate voices on either side of the technical-social validity debate?
4. Do we have the technical capacity, the researchers, the test developers to meet the challenges of the growing expectations that assessments of different kinds will be increasingly central to matters of school reform and other sociopolitical purposes?
5. Should we be advocating national standards and constructing specific assessments to evaluate their achievement? If so, how do we integrate the scoring and interpretation with advanced technologies? What trade-offs between centralized and decentralized approaches to academic standards, to forms of assessment, to norms, to multicultural issues are we, as assessment professionals, prepared to accept?

6. Are we training teachers, counselors, administrators to view assessments in comprehensive terms? Are we helping them understand the political climate for assessment and the validities inherent in using assessment devices?
7. Are we prepared to respond to questions from the media or from policy makers with regard to whether national testing and national standards are good ideas?
8. Within basic education settings, are school counselors major players as resource persons or implementers of assessments for accountability, for exit exams, for other purposes? Are they being trained to play these roles?

While many of the questions that arise from current national and state debates about assessment in education are the content of media headlines, many less publicized but similar questions arise as assessments are being applied to various groups of adolescents and adults: those moving from school to work or welfare to work; dislocated homemakers or women attempting to reenter paid employment; the use of assessments for military applicants in an increasingly technological environment and one which is changing rapidly in proportion of males and females; assessment, including literacy audits, of the competencies of current members of the American work force in their basic academic skills and their teachability or trainability relative to learning new industrial, manufacturing or business processes; the need to insure that persons with disabilities are able to use their talents and skills in educational and work settings without discrimination and bias; the assessment of immigrant populations and cross-national populations being assimilated into or recruited for American jobs for which there are skill shortages.

Embedded in such adult employment initiatives are both implicit and explicit expectations that various types of assessment will be important. And, there are many questions

to be answered. For example, throughout these emphases, there are questions about "What workplace skills are to be assessed?" There is great interest currently in assessing 'softer' skills such as interpersonal skills and work habits, as well as traditional cognitive skills. Other assessment questions proposed at the national level relate to the use of ethnographic approaches to study the use of literacy in the workplace, the importance of informal knowledge for gaining vital on-the-job skills, the use of assessment center procedures developed by AT&T 40 years ago for measuring job skill attributes required for the 21st century, and the use of video or computers to overcome the performance barriers many minority persons experience in taking written tests. Such proposed directions for assessment related to the recruitment, induction, and retraining of persons for the workplace keep pushing out the envelope of available research on such processes and exposes the need for new mindsets and initiatives in assessment (American Educational Research Association, 1996).

Although our time for these remarks does not permit an extended analysis of the issues and challenges that relate to the assessment of these adult populations, suffice it to say that among the assessment issues are the use of standardized, multiple choice, knowledge based tests versus performance-based assessment; new forms of functional analysis for persons with disabilities; sex and racially biased norms; a lack of knowledge about the characteristics and lived experiences of lower socio-economic men and women, their learning styles, their inexperience with assessment processes and how these factors affect their scores; and, certainly as a more pluralistic and culturally diverse population translates into a more culturally diverse work force, how do we create a more diverse assessment system for adults that accommodates language differences, differences in educational backgrounds in the countries of origins from which immigrants are coming and how do we incorporate responses to these issues into tools for

employment and career counselors, rehabilitation counselors, military classification experts, as well as counselors in educational settings.

Given the limitations of time, then, let me turn in the remainder of my remarks to the historic interaction between counseling and assessment, testing and assessment as interventions, and how important assessment has been and is in bridging the gap between theory and practice. Then let me pose some questions about such matters as well.

The alliances between counseling and assessment have ebbed and flowed depending upon what counseling and personality theories were in vogue at particular times in our history, the types of training provided counselors at different points in the past century, and depending on the degree to which assessment has been seen as a legitimate and useful adjunctive input or complement to counseling. A number of the issues of sexual or racial bias of tests among other matters that have already been discussed have affected the use of testing and assessment by counselors.

But, in any case, there are a number of issues affecting assessment in counseling in schools and in other settings which are also hotbed issues and challenges. These issues have to do with a range of process concerns such as are assessment processes being used effectively in counseling, are assessments really interventions in their own right, do assessments effectively bridge the gap between theory and practice, as well as professional issues such as who should test, are counselors and therapists being effectively trained to test, and how do we know that this is true? Let me try to deal with these matters for a few moments in a bit of a potpourri.

While it is tempting to go back to the beginning of the twentieth century and trace the important interaction of counseling and assessment as both have grown in maturity during the century, I will resist that urge. Instead let me suggest briefly that to a large degree changes in

counseling and in assessment have frequently coincided with emerging theories of life span psychological development, client centered or cognitive behavioral counseling, and, particularly, the expanding models of career development of persons like John Holland, Donald Super, John Krumboltz and many others and the attempts of these theorists to make their theories accessible to counselors through the use of assessment instruments. Indeed, given these circumstances, I am frequently puzzled by the continuing criticisms by counselors and by some counselor educators that formal personality, counseling, or career theory is not relevant to what counselors do or that theory and practice are separated because theorists do not tell counselors how to use their theories. I respectfully suggest that in large part that criticism is a myth, rather than a reality. Let me take career development theory as an example. In my view, assessment has been the bridge in operationalizing theoretical constructs by reflecting them in interventions and, in particular, in tests and measurements.

Certainly, this has been true in Holland's theoretical constructs as these are reflected in the Vocational Preference Inventory, My Vocational Situation, and the Self-Directed Search; in the use of his theoretical framework (RIASEC) as the organizing and interpretive structure for the most recent interations of the Strong Interest Inventory and for some of the informational and self-assessment components of the DISCOVER computer-mediated career guidance system; and the use of Holland's three letter coding system of major personality types as a way of organizing U.S. government educational and occupational information through such sources as the Dictionary of Holland codes.

Similarly Super, from the beginning of his conceptual work, has used assessment instruments to operationalize and to evaluate his theoretical constructs. Like Holland, he has made his theoretical constructs accessible to practitioners by using assessment to bridge theory

and practice. Relevant examples would include the Career Development Inventory, the Adult Career Concerns Inventory, the Work Values Inventory, and, more recently, the Values Inventory, the Salience Inventory, and the Career Rainbow. Each of these instruments attempts to describe or to measure individual career behavior in ways that are useful in defining goals for counseling, and in explicating one's maturity or one's levels of career planfulness, knowledge and attitudes about career choice, intrinsic and extrinsic life-career values, and the relative importance to the client of major life roles beyond those of occupation or career.

Super's theoretical work has spawned assessments by others (for example, Crites' Career Maturity Inventory), theoretical extensions like that of Gottfredson's processes of circumscription and compromise as ways of incorporating the effects of gender issues, sex bias and sex roles, as factors shaping the roles of women; and, indeed, models of career counseling, like the recent C-DAC in which assessment and counseling are intimately interactive as interventions.

While there are many other examples, in and out of career theory, of instances in which assessment has been used to bridge theory and practice and, indeed, been conceived as an intervention in its own right, let me finish these notions by briefly acknowledging the importance of John Krumboltz's theoretical concepts through the years, his development of innovative assessment devices during his earlier emphases on behaviorism and as he has articulated his social learning theory and, more recently, his cognitive behavioral theory related to such issues as faulty self-observation generalizations or inaccurate interpretations of environmental conditions and his recent development of the Career Beliefs Inventory as a counseling tool by which to identify presuppositions and irrational beliefs that may block people from achieving their goals.

I must confess that I sometimes wonder whether counselors are being trained to understand the intimacy of theory and the assessment instruments which have been derived from theory and which serve to stimulate client self appraisal as an important intervention alone and as a stimulus to creating the content which counseling explores, clarifies and incorporates in individual plans of action.

But let me not linger on either history or the role of assessment in bridging theory and counseling practice. Rather let me turn quickly to some of the trends of the late 1990s, and some of the challenges for assessment that are spawned by the evolution of counseling in the United States and elsewhere. They include:

- Growing acceptance of counseling programs as central to the mission of schools, higher education institutions and increasingly to workplaces, rather than as frills or ancillary services. In these contexts, assessment and evaluation issues are increasingly seen as major tools pertinent to the integration of institutional missions and the deployment of counseling resources and purposes.
- The systematic development, planning, implementation and evaluation of counseling programs in schools, colleges and universities, and in work places. Such programs are increasingly seen as having their own psychosocial content (e.g. career planning, purposefulness, productivity, stress management, anger reduction) and their own responsibility for facilitating certain types of student or client knowledge, attitudes, and behaviors. Rather than being a random collection of services or functions, or by-products of other activities, counseling programs are increasingly expected to identify the results for which they will be accountable and to provide evidence of their

effectiveness for accountability purposes. Assessment will play an important role both as an intervention within such programs and as an evaluation tool.

- Issues of cost/benefit ratios relative to counseling programs will be increasingly an issue in the next century. In the United States we have rarely raised that issue but our colleagues in Europe have and we can expect that it will be an emerging concern in schools, in workplaces, and in higher education. Assessment strategies will be critical as they relate both to producing relevant measures or assessments of productivity by counselors in different environments and measures of their effectiveness, the outcomes they obtain with individuals or groups of students, clients or consumers as well as the costs of producing the units of productivity measured in relation to, for example, the use of goal directed, time limited interventions, versus psychoeducational models or the use of technology, etc.
- The use of needs assessments related to the topics or problems counselors in different settings should be addressing and the differential treatment by client interactions that should be planned for in designing counseling programs. Again, needs assessment strategies and the measures useful in comparing the effects of differential treatments for specific common outcomes will need to be refined and enhanced.
- Attention to crisis intervention and to addressing the needs of persons at risk (e.g. those who experience chemical dependence or are violence prone, likely to be an academic or work failure, likely to have a teenage pregnancy or to be socially or emotionally dysfunctional) will be a pervasive theme in the next century. Related will be new approaches to early identification, prevention and treatment, more participation of counselors in student assistance programs or employee assistance

programs or other group or shared approaches to intervention for different populations and purposes, and more inclusion of counseling in a total program of interventions aimed at the multiple problems experienced by most people at risk. Assessment needs will pervade such trends. They will include increased emphases on diagnosis, on identifying the competencies needed by counselors with these populations and, in working in combination with other mental health professionals, assessment of differential individual and group treatments, their effectiveness and their cost-benefit ratios.

- We are seeing more uses of technology (e.g. computer assisted career guidance programs, testing and test interpretations by computers, self-directed planning and decision-making, distance learning, electronic information processes) that need to be evaluated both in outcomes and in differential treatment terms.
- In the next century, there is the likelihood of greater use of differentiated staffing among counselors in particular settings and new configurations of counselors, support systems, and technology to deal with different demographic profiles and institutional emphases. Again, there will be increased needs for needs assessment, comparative analyses, cost/benefit research and other types of assessments.
- And there will be an increased emphasis on training counselors in competency-based formats, with different mixes of didactic and hands-on supervision, for work in different settings and with different populations. In these contexts, there will be increased concern about assessing counseling learning styles and preferred modes of training, the use of virtual reality in lieu of or support of the practice of supervised

counseling skills, the types of assessment competencies, as they are applied manually and through technology, that counselors need and within what contexts.

Obviously, such a litany of potential trends is in no way exhaustive, but it suggests what would appear to be a growing need for clarity about how counseling and assessment need to be interactive. Such trends acknowledge that external forces—political, legal, economic, social—will likely modify and/or add to what I have said here about emerging trends describing the importance and character of counseling programs.

One of the growing political and economic challenges for counselors either directly or indirectly is the current national rhetoric about certifying competencies. As the United States continues to engage in school reform, redefinition of workplace education and development, school to work transitions, and workplace reorganization, the nation will place an increasing priority on the certification of competencies possessed by students in schools and universities and by workers. Employers are no longer satisfied to accept program completion as evidence of employability or occupational skill. Instead competency certification at various levels and in different paradigms will be expected and assessment measures will be sought to provide such certification. Given the changing nature of the workplace and in the skills required to work with new industrial and business processes, in technologically intensive environments and in collaborative work groups, one can expect that certification of students or workers competencies will go beyond those of competitiveness, problem-solving ability, resemblance or similarity to work groups, and include greater attention to competencies which underlie complementarity—the ability to facilitate the work of others and engage in group problem-solving;' career motivation—e.g. career resilience, career identity, career insight; personal flexibility and teachability.

The fundamental point here is that the applications of assessments to questions of individual competence and program accountability are going to be major issues far into the 21st Century. And, within such perspectives, counselors are not likely to be exempt from such assessment concerns. The notions of certifying competencies noted above, will extend to counselors as well. Obviously, we are well along that road because of the pioneering leadership in such assessment by AAC, NBCC, NCDA, CACREP and other ACA units. But even given the excellence of these efforts, to date, most of the certification approaches have been knowledge-based, not performance based, at least as they relate directly to the impact of the counselor on clients. These issues are likely to get more delicate in the future as various mental health professional organizations such as APA try to define the scope of practice of their constituents—psychologists—to encompass that which counselors have historically been trained to do and have done. As you may or may not know, counselors in California, Georgia, Indiana, and Louisiana, among other states, have faced recent challenges to fair access to the use of tests in counseling. Psychologists in these states have mounted efforts to restrict the use of tests to doctoral level professionals. The latter is often a code word in specific states for persons who have been trained in APA accredited counseling psychology programs, not in doctoral programs in counselor education. The NBCC Board of Directors has stepped up to this challenge by citing a number of points that it feels are important to the assessment practice of NCCs (Clawson, 1996, 12(3), 1 and 3). They include:

- The practice of counseling requires a right to administer and interpret standardized psychometric assessment instruments (tests) to plan treatment or to assist with life planning.

- The right to administer tests should be based upon, adequate training, not on degree or discipline.
- The current behavior of the American Psychological Association (APA), state psychology licensure boards, and state psychology associations, regarding discipline "ownership" of psychological testing is improper.
- No counselor should administer any assessment instrument without proper training.
- Training institutions should prepare counselors in the proper use and awareness of testing procedures.
- Members of the counseling profession, including all counseling professors, should promote proper use of tests and advocate for counselors' right to use tests.

Let me say that I support everyone of those statements by the NBCC Board but we must recognize that both APA's actions and NBCC's responses are political, not scientific responses. Therefore, at some time in the future, if this issue continues to ferment because of credentialing competition, etc., the resolution will be in answers to more precise assessment questions: who has the competencies that can be demonstrated in their accuracy, relevance, and effectiveness relative to client needs? What is adequate or proper training and how can it be assessed? How do we insure that we recognize the key role in testing of the test user so that we effectively respond to the observation of Anastasie (1992) that: "Most popular criticisms of tests are clearly identifiable as criticisms of test use (or misuse), rather than the tests themselves. Tests are essentially tools. Whether any tool is an instrument of good or bad depends on how the tool is used" (p.610). What are the specific counselor competencies achieved by counselors in training from test and measurement courses as a part of the core preparation for professional counselors? How do these compare with those possessed by psychologists (Masters and Doctoral levels)?

How should counselor competencies differ in relation to the types of tests being used in counseling practice (e.g. the assessment and diagnosis of emotional disorders; the assessment and diagnosis of aptitudes, interests, career maturity, etc.)? These questions and useful answers to them are embedded in ACA's statements of ethics and in the packets of information sent to the Attorneys General of Georgia and Indiana in response to the challenge to counselors' right to do psychological testing in these states. But, these questions and others, that need to be more fully incorporated in counselor competency assessment, are included in such documents as Test User Qualifications: A Data-Based Approach to Promoting Good Test Use (Eyde, Moreland, Robertson, Prinoff & Most, 1988, p. 14) which includes a factor analysis of good testing practices that yielded seven tentative factors and recommendations for fundamental operating principles that include:

1. Comprehensive assessment.	Follow-up to get facts from interpretation psychosocial history to integrate with test scores, as part of interpretation.
2. Proper test use.	Acceptance of responsibility for competent use of the test.
3. Psychometric knowledge.	Consideration of standard error of measurement.
4. Maintaining integrity of test results.	Making clear that cut-off scores imposed for placement in special programs for the gifted are questionable because they disregard measurement error.
5. Accuracy of scoring.	Using checks on scoring accuracy.

6. Appropriate use of norms. Not assuming that a norm for a given job (or group) applies to a different job (or group).

7. Interpretive feedback. Willingness and ability to give interpretation and guidance to test takers, in counseling.

Fundamental operating principles that guided this data-gathering effort were:

1. A model test user qualification system should be based on scientific methods and should serve as a tool for identifying the competencies of test users.
2. Access to psychometric desires should be based on knowledge and behavior of test users rather than solely on job titles or credentials.
3. The key to the model system is self-regulation.
4. The model applies to the broad range of test users who belong to many different professional associations which engage in professional self-regulation, using ethical principles relating to competence.
5. Legislation restricting test use to psychologists or psychologists supervised by psychologists is unrealistic and unnecessarily restrictive, and applies primarily to tests used by psychologists, thus ignoring other practitioners.
6. Test misuse is more likely to be a function of lack of information or misinformation than of malfeasance on the part of the test user.
7. Educational efforts, rather than restriction of access, are likely to be most effective in promoting good testing practices.
8. The proposed competency-based model user qualifications system, which is designed to reduce test misuse, is likely to increase the use of tests as an important element in decision making.

9. By identifying possible test misuse, the system will alert test users to poor testing practices and reduce the likelihood that tests will be banned through legislative action.

This document was derived from an interdisciplinary model that represents an alliance of organizations whose practitioners engage in assessment; it is an important and enlightened approach to the credentialing mania and turfdom which is again arising. Interestingly, the operating principles embodied in this document are essentially those that are included in the ACA Ethical Standards (as revised, April, 1995) and they are consistent with the counselor preparation standards, etc., promulgated by CACREP.

Unfortunately, as I have said earlier, what each of us must recognize is that this growing assault on counselors' use of assessment by psychologists is not a matter of science or of aggregated research findings but instead of power, protection of the independent marketplace, and politics.

As the issues of power and politics continue to arise about who should be able to test and within what scope of work, one of the concepts that helps to explain the stimulus to such tensions among professional groups is the fact that tests, for reasons we have already addressed, have become terribly important elements in contemporary society. They are important in the accuracy or inaccuracy of their content in relation to their purposes, they are important in their application and interpretation, they are important in their uses for classification, for inclusion or exclusion, and they are important in the populations for which they are relevant. They are also world wide in their development and in their use. Indeed, there are a growing number of recent conferences that have been held in Greece, Spain, Germany, and Belgium that have focused on the use of psychological assessment. Recently, Division 2 of the International Association of Applied

Psychology has begun the process of developing "Guidelines for Adopting Psychological Tests for Use in Multiple Languages and Cultures." Underlying such notions are the increasing use of tests cross-culturally and the validity of such uses. Dale Prediger's (March, 1993) editing of Multicultural Assessment Standards: A Compilation for Counselors is an excellent reference which puts many of these issues into context.

While there is much more to be said about the challenges just cited, let me turn in conclusion to some final assessment issues with which I will deal only briefly.

(1) One is teaching the test. Certainly the use of tests as diagnostic instruments to identify developmental deficits or psychological traits or states of different forms of maximum behavior as in aptitude tests or typical behavior as found in attitudes or interests have a long and important history. However, often we consider all of the scores from these assessments as fixed effects rather than as fixed effects in some cognitive or behavioral areas but in other areas measures of more malleable individual characteristics that are susceptible to learning on the part of the individual, particularly as it can be guided by teaching persons why their answers to the tests they took were wrong and what is implied for them in learning or relearning certain types of behavior or knowledge. In such cases, depending upon their uses, tests can be interventions. This is a different mentality about testing than suggesting that the scores attained are absolute scores and therefore not susceptible to modification. One can argue that coaching people to do better in the SAT, the MCAT or various occupational entrance examinations is teaching the test and that the data are mixed about whether such coaching does any good. True enough! But part of the response is that teaching the test depends upon what tests you are talking about (e.g. career

development process instruments, etc.) and how you view related observations by Charles Healy and others that focus on helping clients to develop self-assessment skills and to being true collaborators in the appraisal process. In particular, Healy has talked about four obstacles to such a shift in thinking about new counselor-client collaborative models of using appraisal data in career counseling. We have all heard them many times. They include: "(1) casting clients as subordinates rather than as collaborators; (2) discounting self-assessment by favoring counselor assessments; (3) de-emphasizing the influence of contexts in client development; and (4) focusing on a single choice rather than on strengthening client decision making and knowledge for follow-through" (p.214). Such views are obviously impediments to the growing need to empower the client by giving him or her the assessment skills that can be strengthening by teaching the test and by using the content to encourage client self-evaluation and decision making in ways, as Prediger (1994), Zytowski (1994), and Kapes, Mastie, and Whitfield (1994), in their superb book, have also discussed by which the relationship between testing and counseling can be enhanced, not fragmented.

(2) A further challenge, although not necessarily a new one, has to do with computer applications to testing. While in one sense such applications have become commonplace, such applications are also uneven in their use across settings, populations, geographical regions, etc. Computer applications to testing include the self assessments embedded in computer assisted career guidance programs but they go beyond such applications to the administration, scoring and interpretation of tests. We are finding increased use of computers for self-help programs of all kinds

> including those purporting to provide personal counseling, as expert systems modeling counselor behavior in responding to a client's descriptions of his or her psychological dilemmas to the computer. The computer is also being used by some health personnel for consultation on emotional crises where psychiatrists, psychologists or counselors are not immediately available.

The fundamental point is that the use of computers in testing, in statistical analysis, and in all sorts of other related ways sometimes occur in immediate conjunction to the process of counseling or psychotherapy not simply as administrative procedures unconnected to counseling. As a result, there are continuing and in some ways increasingly complex ethical questions involved in the application of computers to testing, to self-appraisal, to personal counseling and to the variations on these themes.

Computer-based test administration and interpretation, like every other technique available to a counselor, can be both a boon and a bane. On the positive side, they can be cost-effective and, in the case of microcomputers, provide test information virtually instantaneously. In general, clients seem to enjoy the experience and to achieve as much self-knowledge as when paper-and-pencil tests are used. Further, no violence seems to be visited on the psychometric properties of accepted testing instruments that are computerized (that is, validity, reliability, and so on). The negatives of computer-based testing are more involved with the idiosyncratic aspects of a particular instrument, interpretive programs, or hardware configurations than with the idea itself. Group administration is obviously difficult, if not impossible, because of the prohibitive cost of multiple stations; some programs are not user-friendly; some instruments are so new and rushed to market so quickly that they provide inadequate validity and normative data; and erroneous or overly generalized interpretations are possible. Further, I will not take time to

cite these but to note the reality that there is as yet little research to determine individual differences in person-machine interactions. A final concern is perhaps the most ominous, counselors may believe that because the machine is producing an impressive-looking report, they need not have an in-depth knowledge of the test, its underlying constructs, its psychometric strengths and weaknesses, appropriate interpretations, and the need to integrate the results with everything else of relevant importance in the career development of the client.

I would be remiss here if I did not mention the challenges to assessment and to the ethics of assessment that are now inherent in the Internet. I do not have to remind you that we have, as a nation, embraced the Internet with a passion that belies the reality that there has been virtually no research done about the effects of the Internet on learning, mental health, career decision-making, etc. Some 30 to 50,000 pages are being added to the Internet each day and much of this content purports to be relevant to what counselors do.

Inherent in the Internet is concern about ethical research in the information age. The implication is that researchers who study electronic communities or on-line communities will likely find themselves increasingly using qualitative methods, changing their commonly used research tools, and adapting their assessment strategies to these new electronic environments. In essence, each of the current capabilities of the Internet from e-mail to chat rooms will pose its own research and assessment dilemmas related to how to obtain informed consent; how to conceive respondents as owners of the materials they create; how to protect copywritten material on the net; how to create a climate of trust, collaboration and equality with electronic community members; how to negotiate researcher entrée into an electronic community; how to treat electronic mail as private correspondence, not be treated as research data unless express permission is given; how to respect the identity of the research respondents in an electronic

community, protect or mask the origins of the communications and communicate the results of their research to participants in the research? (Schrum, 1995)

In conclusion, then, let me say what you already know—there is much more to say and challenges and issues that are yet to be identified that relate to assessment in the 21st Century. Many of these issues and challenges will be identified, discussed, and explored in other speeches and content sessions in this conference. But, what is apparent in these deliberations is the reality that during the 20th Century both assessment and counseling have sunk their roots deep into the American social fabric and both have matured in their conceptual and methodological processes. Both will be extremely important in the 21st Century as they contribute to national goals of mental health, career development, productivity and individual purpose. But to do so continues to raise questions that must be addressed systematically and scientifically. They include:

1. Are we training counselors in the most effective ways to use assessments, to understand their roles as interventions, and as integral to counseling processes?
2. Have we identified effective training models in counselor use of assessments and the competencies necessary to use different types of tests in counseling practice? Are we providing sufficient opportunities to retrain counselors in assessment whose skills and understanding of assessment may be outdated?
3. Are we training counselors to use assessments in new and emerging contexts: teach the test, use computer assisted test interpretations, use the World Wide Web to do assessments?
4. Have we considered how different groups of helping professionals differ in their ability to use assessments and how they might complement each other in school, community, or workplace contexts?

5. Are we as specialists or users of assessments prepared to understand the political as well as the scientific and technical issues related to the uses of assessment in counseling?
6. Are we prepared to talk about the cost-benefit ratios of assessment, used in different forms and models, and in relation to different models of counseling practice? Are we prepared to talk about the assessment of counseling both in terms of productivity and of effectiveness?
7. Are we adequately preparing counselors to think and act in multi-cultural terms as they address assessment issues? Do counselors understand and act in accord with existing research that shows that persons from different national and cultural traditions, even if they are residents of a pluralistic nation like the United States, may have different values, beliefs, communications styles and methods of solving problems, perceive problems differently and cope with them in different ways (Wilgosh & Gibson, 1994).

While not exhaustive of the questions before us, they like those earlier inventoried in these remarks, are representative of the issues and challenges that await us in 1998 and beyond.

Thank you for your attention.

References

American Counseling Association. (1995, April). Code of ethics and standards of practice. Alexandria, VA: The Author.

American Educational Research Association. (1996, March). Research Policy Notes.

Anastadsie, A. (1992). What counselors should know about the use and interpretation of psychological tests. Journal of Counseling & Development, 70(5), 610-615.

Archer, J. (1997, October 29). New test in r.1. sends shock waves through state. Education Week, 15.

Clawson, T. (1996, Winter). Fair access in testing. NBCC News Notes, 12(3), 1, 3.

Gordon, E.W., & Terrell, M.D. (1981). The changed social context of testing. American Psychologist, 36 (10), 1167-1171.

Eyde, L.D., Moreland, K.L., Robertson, G.J., Primoff, E.S., & Most, R.B. (1988, December). Test user qualifications: A data-based approach to promoting good test use. Issues in scientific psychology. Washington, D.C.: American Psychological Association, Science Directorate.

Hoff, J. (1997, November 5). Latest testing compromise appears doomed. Education Week, 22, 26.

Johnson, R.C. (1997, October 22). Michigan House passes bill be revise high school testing program. Education Week, 11.

Johnson, R.C. (1997, October 29). Education prominent in N.J. governor's race. Education Week, 15, 17.

Kapes, J.T., Mastie, M.M., & Whitfield, E.A. (1994). A counselor's guide to career assessment instruments (3rd Ed.). Alexandria, VA: National Career Development Association.

Lawton, M. (1997, October 22). Discrimination claimed in Texas exit-exam lawsuit. Education Week, 3.

Lawton, M. (1997, October 22). State board leaders call for assessments bearing consequences. Education Week, 7.

Manzo, K.K. (1997, October 22). High stakes: Test truths or consequences. Education Week, 1, 9.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performance as scientific inquiry into score meaning. American Psychologist, 50, 741-749.

Prediger, D.J. (Ed.). (1993, March). Multicultural assessment standards for counselors. Alexandria, VA: the Association for Assessment in Counseling.

Reinhard, B. (1997, October 22). Reporter's notebook. Education Week, 5.

Sack, J. (1997, October 29). Proposed IDEA rules target testing, accountability. Education Week, 19, 24.

Schrum, L. (1995). Framing the debate: Ethical research in the information age. Qualitative Inquiry, 1(3), 311-326.

Supovitz, J.A. (1997, November 5). From multiple choice to multiple choices. A diverse society deserves a more diverse assessment system. Education Week, 34, 37.

Terwilliger, J. (1997). Semanties, psychometrics, and assessment reform: A close look at 'authentic' assessments. Educational Researcher, 26(8), 24-27.

White, K.A.. (1997, November 5). Duty-bound. Education Week, 28-32.

Wilgosh, L., & Gibson, J.T. (1994). Gross-national perspectives on the role of assessment in counselling: A preliminary report. International Journal for the Advancement of Counselling. 17, 59-70.

Zytowski, D.G. (1994). Tests and counseling: We are still married and living in discriminant analysis. Measurement and Evaluation of Counseling and Development, 26, 219-223.